GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: Mar... 11, 20..2, 17:0...:29 ; Search time 1725.6 Seconds
(without alignments)
650.097 Million cell updates/sec

Title: US-09-786-009-3
Perfect score: 68
Sequence: 1 tcaaatctaaacatatqqc.....ctqcatcacqqaaqatqca 68

Scoring table: IDENTITY_NUC
Gapop 1.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters. 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vi:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rod:*
33: em_htg_mam:*
34: em_htg_inv:*
35: em_htg_rod:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	28.2	41.5	1e+127	2	AC025779	AC025779 Homo sapi
c 2	28.2	41.5	1e+127	2	AC018883	AC018883 Homo sapi
3	28.2	41.5	243615	2	AC092333	AC092333 Homo sapi
c 4	28.2	41.5	2e+008	2	AC092362	AC092362 Homo sapi
5	28	41.2	257	7	AC007377	AC007377 Xenopus l
6	27.8	40.9	27500	2	AC011617	AC011617 Leishmani
c 7	27.8	40.9	1e+008	2	AC012418	AC012418 Homo sapi
c 8	27.8	40.9	1.85e+	2	AC [illegible]	AC [illegible] Hom sapi
c 9	27.8	40.9	166178	2	AC023060	AC023060 Homo sapi
10	27.8	40.9	170643	9	AL158421	AL158421 Human DNA
11	27.8	40.9	261741	2	AC [illegible]	AC [illegible] Hom sapi
c 12	27.2	40.0	2740	1	D67517	D67517 Methanococ
13	27.2	40.0	4227	1	D67472	D67472 Methanococ
c 14	27	39.7	2605	6	AB [illegible]	AB [illegible] Sequencing
c 15	26.6	39.1	137137	9	AC004928	AC004928 Homo sapi
16	26.2	38.5	12627	2	AC087738	AC087738 Pan troglo
c 17	26.2	38.5	2e+006	2	AC [illegible]	AC [illegible] Hom sapi
18	26	38.2	2465	1	AJ276242	AJ276242 Sus scrof
19	26	38.2	3242	4	BX279612	AJ276242 Sus scrof
20	26	38.2	15254	9	BC031210	Y14214 Homo sapien
21	26	38.2	17480	9	BC [illegible]	Y14214 Homo sapien
22	26	38.2	117483	2	AC [illegible]	AC [illegible] Hom sapi
23	26	38.2	1.09e+	2	AC [illegible]	AC [illegible] Hom sapi
24	26	38.2	136774	2	AC [illegible]	AC [illegible] Hom sapi
c 25	25.8	37.9	154e+	2	AC [illegible]	AC [illegible] Hom sapi
c 26	25.8	37.9	175888	2	AC [illegible]	AC [illegible] Hom sapi
c 27	25.8	37.9	2.34e+	2	AC [illegible]	AC [illegible] Hom sapi
28	25.4	37.4	1.99e+	9	AL [illegible]	AL [illegible] Human DNA
29	25.4	37.4	134882	2	AC [illegible]	AC [illegible] Hom sapi
30	25.4	37.4	210842	2	AC [illegible]	AC [illegible] Mus musc
c 31	25.2	37.1	199e+	9	AK024606	AK024606 Homo sapi
c 32	25.2	37.1	102435	8	AY014124	AY014124 Homo sapi
c 33	25.2	37.1	150829	2	AC [illegible]	AC [illegible] Hom sapi
c 34	25.2	37.1	151865	2	AL356239	AL356236 Homo sapi
35	25.2	37.1	192373	2	AP003551	AP003551 Oryza sat
36	25.2	37.1	165120	9	AL138704	AL138704 Human DNA
c 37	25.2	37.1	167422	2	AC024903	AC024903 Homo sapi
c 38	25.2	37.1	168951	2	AC016927	AC016927 Homo sapi
c 39	25.2	37.1	172256	2	AC [illegible]	AC [illegible] Hom sapi
40	25.2	37.1	178806	2	AC013557	AC013557 Homo sapi
41	25.2	37.1	191616	9	AC009162	AC009162 Homo sapi
c 42	25	36.8	1225	8	HVAJ4e7	AC004e57 Hordeum v
c 43	25	36.8	49838	9	HS387E22	AL138680 Human DNA
44	25	36.8	110952	2	AP004412	AP004 42 Oryza sat
c 45	25	36.8	136344	2	AL590 [illegible]	AC [illegible] Hom sapi

AC NM:N

RESULT 1
AC025779/c
LOCUS AC025779 181127 bp DNA HTG 01-JUL-2
DEFINITION Homo sapiens chromosome 5, clo... ...ete ... ...ce. W BRING [illegible]
SEQUENCE, * ordered ...ces.
ACCESSION AC025779
VERSION AC025779.5 GI:[illegible]
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; ...AFT; VERTU.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 181127)
AUTHORS DOE Joint Genome Institute
TITLE Sequencing of Human Chromosome
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 181127)
AUTHORS DOE Joint Genome Institute
TITLE Direct Submission
JOURNAL Submitted (14-MAR-2... by DOE ... ...e, ... Facility, ...

```
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Apr 19, 2001 this sequence version replaced gi:8576152.
                   Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov/ [illegible]

            Project Information
            Center Project Name: 558008
            Center clone name: RPCI 11 374E21

            Summary Statistics
            Consensus quality: 176377 bases at least Q40
            Consensus quality: 179054 bases at least Q30
            Consensus quality: 180066 bases at least Q20
            Estimated insert size: 188470; agarose fp estimation
            Estimated insert size: 180627; sum of contigs estimation
            Quality coverage: 7.32 in Q20 bases; agarose fp estimation
            Quality coverage: 7.64 in Q20 bases; sum of contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 6 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *          1     28393: contig of 28393 bp in length
            *      28394     28493: gap of unknown length
            *      28494     29162: contig of 669 bp in length
            *      29163     29262: gap of unknown length
            *      29263    105192: contig of 75930 bp in length
            *     105193    105292: gap of unknown length
            *     105293    111944: contig of 6652 bp in length
            *     111945    112044: gap of unknown length
            *     112045    171230: contig of 59186 bp in length
            *     171231    171330: gap of unknown length
            *     171331    181127: contig of 9797 bp in length.
FEATURES             Location/Qualifiers
     source          1..181127
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RPCI 11-374E21"
                     /clone_lib="RPCI human BAC library 11"
BASE COUNT   [illegible] others
ORIGIN

  Query Match             41.5%;  Score 28.2;  DB 2;  Length 181127;
  Best Local Similarity   68.4%;  Pred. No. 33;
  Matches   39;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

  [illegible alignment lines]

RESULT 2
AC018883/c
LOCUS       AC018883   197523 bp    DNA             HTG       07-JUL-2001
DEFINITION  Homo sapiens chromosome 5 clone RP11-36012, WORKING DRAFT SEQUENCE,
            18 unordered pieces.
ACCESSION   AC018883
VERSION     AC018883.4  GI:9568169
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 197523)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 197523)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (21 DEC 1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jun 16, 2000 this sequence version replaced gi:6850700.

                   Genome Center
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site: http://[illegible]/index.shtml
            ------         Project Information         ------
            Center project name: H_NH0360I02
            ------         Summary Statistics          ------
            Sequencing vector: M13; 87%
            Sequencing vector: plasmid; 13%
            Chemistry: Dye-primer ET; 87% of reads
            Chemistry: Dye-terminator Big Dye; 13% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 186448 bases at least Q40
            Consensus quality: 189098 bases at least Q30
            Consensus quality: 192268 bases at least Q20
            Insert size: 190000; agarose-fp
            Insert size: 195824; sum-of-contigs
            Quality coverage: 4.25 in Q20 bases; agarose-fp
            Quality coverage: 4.37 in Q20 bases; sum-of-contigs
            ----------------------------------------------------
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 18 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *          1      2637: contig of 2637 bp in length
            *       2638      2737: gap of unknown length
            *       2738      4614: contig of 1877 bp in length
            *       4615      4714: gap of unknown length
            *       4715      6872: contig of 2158 bp in length
            *       6873      6972: gap of unknown length
            *       6973     10123: contig of 3151 bp in length
            *      10124     10223: gap of unknown length
            *      10224     14303: contig of 4080 bp in length
            *      14304     14403: gap of unknown length
            *      14404     18798: contig of 4395 bp in length
            *      18799     18898: gap of unknown length
            *      18899     24681: contig of 5783 bp in length
            *      24682     24781: gap of unknown length
            *      24782     30335: contig of 5554 bp in length
            *      30336     30435: gap of unknown length
            *      30436     36923: contig of 6488 bp in length
            *      36924     37023: gap of unknown length
            *      37024     46432: contig of 9409 bp in length
            *      46433     46532: gap of unknown length
            *      46533     55361: contig of 8829 bp in length
            *      55362     55461: gap of unknown length
            *      55462     67835: contig of 12374 bp in length
            *      67836     67935: gap of unknown length
            *      67936     84414: contig of 16479 bp in length
            *      84415     84514: gap of unknown length
            *      84515    100860: contig of 16346 bp in length
            *     100861    100960: gap of unknown length
            *     100961    121449: contig of 20489 bp in length
            *     121450    121549: gap of unknown length
            *     121550    140877: contig of 19328 bp in length
            *     140878    140977: gap of unknown length
            *     140978    160159: contig of 19182 bp in length
            *     160160    160259: gap of unknown length
            *     160260    197523: contig of 37264 bp in length.
```

```
FEATURES             Location/Qualifiers
     source          1..197523
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RP11-336O12"
     misc_feature    1..2857
                     /note="assembly_name:Contig7"
     misc_feature    2738..4614
                     /note="assembly_name:Contig8"
     misc_feature    4715..6872
                     /note="assembly_name:Contig9"
     misc_feature    6973..10123
                     /note="assembly_name:Contig10"
     misc_feature    10224..14403
                     /note="assembly_name:Contig11"
     misc_feature    14404..18798
                     /note="assembly_name:Contig12"
     misc_feature    18899..24681
                     /note="assembly_name:Contig13"
     misc_feature    24782..30435
                     /note="assembly_name:Contig14"
     misc_feature    30436..36923
                     /note="assembly_name:Contig15"
     misc_feature    37024..46432
                     /note="assembly_name:Contig16"
     misc_feature    46533..55361
                     /note="assembly_name:Contig17"
     misc_feature    55462..67845
                     /note="assembly_name:Contig18"
     misc_feature    67946..84414
                     /note="assembly_name:Contig19"
     misc_feature    84515..100860
                     /note="assembly_name:Contig20"
     misc_feature    100961..121449
                     /note="assembly_name:Contig21"
     misc_feature    121550..140877
                     /note="assembly_name:Contig22"
     misc_feature    140978..160159
                     /note="assembly_name:Contig23
                     clone_end:T7
                     vector_side:left"
     misc_feature    160260..197523
                     /note="assembly_name:Contig24"
BASE COUNT    63017 a   34606 c   34389 g   63803 t    1708 others
ORIGIN


  Query Match             41.5%;  Score 28.2;  DB 2;  Length 197523;
  Best Local Similarity   68.4%;  Pred. No. 32;
  Matches   39;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 tcaaatctaaacatatgcaatatgaatgacggcagcaactgaaggctattccctgcatca 57
           | | |||| |       ||| ||| |||   | | ||| || | |||||||||| |||
Db 190675 TGGCATCTTGCTATGTTGCCTTGGCTGCACTCAGGCTCAAGAGATCTTCCTGCCTCA 190619


RESULT   3
AC092333
LOCUS       AC092333    242015 bp    DNA             HTG       03-JUL-2001
DEFINITION  Homo sapiens chromosome 5 clone RP11-196A23, WORKING DRAFT
            SEQUENCE, 28 unordered pieces.
ACCESSION   AC092333
VERSION     AC092333.1  GI:14589522
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      human
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 242015)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished

REFERENCE   2  (bases 1 to 242015)
  AUTHORS   DOE Joint Genome Institute
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Dr., Walnut Creek, CA 94598, USA
COMMENT           Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 336O12
            Center clone name: RP11-196A23
            -------
            Summary Statistics
            Consensus quality: 99.91% bases > low_q 40
            Consensus quality: 99.62% bases > low_q 50
            Consensus quality: 99.02% bases > low_q 60
            Estimated insert size: 158.1 +/- 30.9 by 3rd party estimator
            Estimated insert size: 237.1 +/- 50.7 by this estimator
            Quality coverage: 5.4 in 120 bases - standard estimator
            Quality coverage: 5.5 in 120 bases - improved contigs estimator
            * NOTE: This is a working draft sequence, currently
            * consists of 48 contigs. The order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the pieces are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1..1074       contig of 1 seq. 1074 bp, 0 low qual
            *    1075..1174       gap of unknown length
            *    1175..2445       contig of 1 seq. 1271 bp, 0 low qual
            *    2446..2479       gap of unknown length
            *    2480..3580       contig of 1 seq. 1101 bp, 0 low qual
            *    3581..4967       gap of unknown length
            *    4968..6592       contig of 1 seq. 1625 bp, 0 low qual
            *    6593..6602       gap of unknown length
            *    6603..7839       contig of 1 seq. 1237 bp, 0 low qual
            *    7840..7939       gap of unknown length
            *    7940..9290       contig of 1 seq. 1351 bp, 0 low qual
            *    9291..10098      gap of unknown length
            *   10099..11098      contig of 1 seq. 1000 bp, 0 low qual
            *   11099..13106      gap of unknown length
            *   13107..14206      contig of 1 seq. 1100 bp, 0 low qual
            *   14207..16139      gap of unknown length
            *   16140..16239      contig of 1 seq. 100 bp, 0 low qual
            *   16240..17477      gap of unknown length
            *   17478..17578      contig of 1 seq. 101 bp, 0 low qual
            *   17579..18993      gap of unknown length
            *   18994..19093      contig of 1 seq. 100 bp, 0 low qual
            *   19094..20665      gap of unknown length
            *   20666..20765      contig of 1 seq. 100 bp, 0 low qual
            *   20766..23023      gap of unknown length
            *   23024..23423      contig of 1 seq. 400 bp, 0 low qual
            *   23424..24951      gap of unknown length
            *   24952..25051      contig of 1 seq. 100 bp, 0 low qual
            *   25052..27411      gap of unknown length
            *   27412..27511      contig of 1 seq. 100 bp, 0 low qual
            *   27512..30588      gap of unknown length
            *   30589..30688      contig of 1 seq. 100 bp, 0 low qual
            *   30689..32741      gap of unknown length
            *   32742..32841      contig of 1 seq. 100 bp, 0 low qual
            *   32842..34701      gap of unknown length
            *   34702..34801      contig of 1 seq. 100 bp, 0 low qual
            *   34802..38296      gap of unknown length
            *   38297..38396      contig of 1 seq. 100 bp, 0 low qual
            *   38397..41546      gap of unknown length
            *   41547..41646      contig of 1 seq. 100 bp, 0 low qual
            *   41647..44431      gap of unknown length
```

```
*     44458     44477: gap of unknown length
*     44478     46461: contig of 2044 bp in length
*     46462     46561: gap of unknown length
*     46562     48475: contig of 2914 bp in length
*     48476     49575: gap of unknown length
*     49576     51675: contig of 2100 bp in length
*     51676     52075: gap of unknown length
*     52076     54564: contig of 2489 bp in length
*     54565     54664: gap of unknown length
*     54665     57495: contig of 2831 bp in length
*     57496     57595: gap of unknown length
*     57596     65813: contig of 8218 bp in length
*     65814     65913: gap of unknown length
*     65914     72608: contig of 6695 bp in length
*     72609     72708: gap of unknown length
*     72709     77127: contig of 4419 bp in length
*     77128     77227: gap of unknown length
*     77228     83304: contig of 6077 bp in length
*     83305     83404: gap of unknown length
*     83405     88866: contig of 5462 bp in length
*     88867     88966: gap of unknown length
*     88967     97138: contig of 8172 bp in length
*     97139     97238: gap of unknown length
*     97239    102438: contig of 5200 bp in length
*    102439    102538: gap of unknown length
*    102539    112046: contig of 9508 bp in length
*    112047    112146: gap of unknown length
*    112147    120360: contig of 8214 bp in length
*    120361    120460: gap of unknown length
*    120461    128470: contig of 8010 bp in length
*    128471    128570: gap of unknown length
*    128571    136687: contig of 8117 bp in length
*    136688    136787: gap of unknown length
*    136788    144774: contig of 7987 bp in length
*    144775    144874: gap of unknown length
*    144875    153436: contig of 8562 bp in length
*    153437    153642: gap of unknown length
*    153643    164737: contig of 11095 bp in length
*    164738    164837: gap of unknown length
*    164838    174879: contig of 10042 bp in length
*    174880    174979: gap of unknown length
*    174980    184262: contig of 9283 bp in length
*    184263    184662: gap of unknown length
*    184663    195668: contig of 11006 bp in length
*    195669    195868: gap of unknown length
*    195869    205038: contig of 9170 bp in length
*    205039    205138: gap of unknown length
*    205139    217691: contig of 12553 bp in length
*    217692    217791: gap of unknown length
*    217792    229073: contig of 11282 bp in length
*    229074    229273: gap of unknown length
*    229274    241615: contig of 12342 bp in length.
FEATURES             Location/Qualifiers
     source          1..241615
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RPCI 11-392A20"
                     /clone_lib="RPCI human BAC library 11"
BASE COUNT     75287 a   43411 c   43607 g   75273 t    4881 others
ORIGIN


  Query Match             41.7%;  Score 28.2;  DB 2;  Length 241615;
  Best Local Similarity   68.4%;  Pred. No. 81;
  Matches   39;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 ttctcatctaaaccatatcaatgaatgaccgcctcgtcgaggtcttcctcgcatcca 57
           || || || | || | || | | | ||| || | | |||||||| |||
Db  86611 ttgtatctgtgatatctgtatatcttcatgtcgacctcagggtcaacgatcatc 86667


RESULT    4
```

```
ACCESSION
LOCUS       AC092362    262608 bp    DNA              HTG       03-JUL-2001
DEFINITION  Homo sapiens chromosome 5 clone RPCI-392P13, WORKING DRAFT
            SEQUENCE, 29 unordered pieces.
ACCESSION   AC092362
VERSION     AC092362.1  GI:14[illegible]
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 262608)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 262608)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JUL-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     --    - Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://[illegible]
            --
            Project Information
            Center Project Name: 7266
            Center clone name: RPCI 11-392P13
            --
            Summary Statistics
            Consensus quality: 238118 bases at least Q40
            Consensus quality: 252698 bases at least Q30
            Consensus quality: 254471 bases at least Q20
            Estimated insert size: 167930; agarose-fp estimation
            Estimated insert size: 259808; sum-of-contigs estimation
            Quality coverage:  14.99 in Q20 bases; agarose-fp estimation
            Quality coverage:  9.67 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 29 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1473: contig of 1473 bp in length
            *     1474     1573: gap of unknown length
            *     1574     2720: contig of 1147 bp in length
            *     2721     2820: gap of unknown length
            *     2821     4142: contig of 1322 bp in length
            *     4143     4242: gap of unknown length
            *     4243     5814: contig of 1572 bp in length
            *     5815     5914: gap of unknown length
            *     5915     6987: contig of 1073 bp in length
            *     6988     7087: gap of unknown length
            *     7088     8448: contig of 1361 bp in length
            *     8449     8548: gap of unknown length
            *     8549     9583: contig of 1035 bp in length
            *     9584     9683: gap of unknown length
            *     9684    11798: contig of 2115 bp in length
            *    11799    11808: gap of unknown length
            *    11809    14149: contig of 2241 bp in length
            *    14150    14239: gap of unknown length
            *    14240    16720: contig of 2481 bp in length.
            *    16721    16820: gap of unknown length
            *    16821    23149: contig of 6329 bp in length
            *    23150    23249: gap of unknown length
            *    23250    29355: contig of 6106 bp in length
            *    29356    29455: gap of unknown length
            *    29456    35681: contig of 6226 bp in length
            *    35682    35781: gap of unknown length
            *    35782    42019: contig of 6238 bp in length
            *    42020    42119: gap of unknown length
```

```
*     42120     49804: contig of 7685 bp in length
*     49805     49904: gap of unknown length
*     49905     58842: contig of 8938 bp in length
*     58843     58942: gap of unknown length
*     58943     68221: contig of 9279 bp in length
*     68222     68321: gap of unknown length
*     68322     78413: contig of 10092 bp in length
*     78414     78513: gap of unknown length
*     78514     90243: contig of 11730 bp in length
*     90244     90343: gap of unknown length
*     90344    105334: contig of 14991 bp in length
*    105335    105434: gap of unknown length
*    105435    115033: contig of 9599 bp in length
*    115034    115133: gap of unknown length
*    115134    122391: contig of 7258 bp in length
*    122392    122491: gap of unknown length
*    122492    129865: contig of 7374 bp in length
*    129866    129965: gap of unknown length
*    129966    144625: contig of 14660 bp in length
*    144626    144725: gap of unknown length
*    144726    165714: contig of 20989 bp in length
*    165715    165814: gap of unknown length
*    165815    186267: contig of 20453 bp in length
*    186268    186367: gap of unknown length
*    186368    199645: contig of 13278 bp in length
*    199646    199745: gap of unknown length
*    199746    219844: contig of 20099 bp in length
*    219845    219944: gap of unknown length
*    219945    262608: contig of 42664 bp in length.
FEATURES             Location/Qualifiers
     source          1..262608
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="RP11-492F13"
                     /clone_lib="RPCI human BAC library 11"
BASE COUNT    83194 a  47141 c  49923 g  82192 t   2855 others
ORIGIN

  Query Match             11.5%;  Score 28.2;  DB 2;  Length 262608;
  Best Local Similarity   68.4%;  Pred. No. 30;
  Matches   39;  Conservative    0;  Mismatches   18;  Indels    0;  Caps    0;

Qy          1 tcgaatctaaacatatagccatgaggtaaggccacctgaaggtcttcctgcatca 57
              | | |||| |  || | | |.| ||| |||    | | ||| |   | |||||| || |||
Db      32933 TGGCATCTTGCTATGTTGCCTTGGCTGGACTGAGGTCAAGAGATCTTCCTGCCTCA 32877

RESULT 5
XLAJ9313
LOCUS       XLAJ9313      857 bp    mRNA            VRT       28-JUL-1999
DEFINITION  Xenopus laevis cDNA clone 32B12 1.
ACCESSION   AJ009313
VERSION     AJ009313.1  GI:3821790
KEYWORDS    .
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 857)
  AUTHORS   Gawantka,V., Pollet,N., Delius,H., Vingron,M., Pfister,R.,
            Nitsch,R., Blumenstock,C. and Niehrs,C.
  TITLE     Gene expression screening in Xenopus identifies molecular pathways
            predicts gene function and provides a global view of embryonic
            patterning
  JOURNAL   Mech. Dev. 77 (2), 95-141 (1998)
  MEDLINE   99051185
REFERENCE   2  (bases 1 to 857)
  AUTHORS   Nicolas,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JUN-1998) Department of Molecular Embryology,
            Deutsches Krebsforschungszentrum [illegible] Neuenheimer Feld 280,
            Heidelberg D-69120, Germany
FEATURES             Location/Qualifiers
     source          1..857
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
BASE COUNT      184 a    211 c    227 g    [illegible]
ORIGIN

  Query Match             11.2%;  Score 28.[illegible];  [illegible]  Length 857;
  Best Local Similarity   66.7%;  Pred. No. [illegible];
  Matches   48;  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          [illegible] tctaaacatatagccatgaggtcc[illegible]gcatcaggaggaat [illegible]
              [illegible]
Db        562 TCTACATGAT[illegible]AAGAAGGAAGAT [illegible]

RESULT 6
AC011617
LOCUS       AC011617     37500 bp    DNA             HTG       29-APR-2[illegible]
DEFINITION  Leishmania major chromosome 35 [illegible] strain Friedlin, ***
            SEQUENCING IN PROGRESS ***, [illegible] ordered pieces.
ACCESSION   AC011617
VERSION     AC011617.2  GI:7670[illegible]
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Leishmania major.
  ORGANISM  Leishmania major
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Leishmania.
REFERENCE   1  (bases 1 to 37500)
  AUTHORS   Myler,P.J., [illegible]
            McDonagh,P. and Stuart,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-OCT-19[illegible]) Seattle [illegible] Research Institute, [illegible]
            Nickerson Street, Seattle, WA 98[illegible] USA
COMMENT     On Apr 29, 2000 this sequence version replaced gi:6[illegible]6666.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submitter.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *          1     28345: contig of 28345 bp in length
            *      28346     28[illegible]: gap of unknown length
            *      28[illegible]     33845: contig of [illegible]4 bp in length
            *      33846     338[illegible]: gap of unknown length
            *      33871     37500: contig of [illegible] bp in length.
FEATURES             Location/Qualifiers
     source          1..37500
                     /organism="Leishmania major"
                     /strain="Friedlin"
                     /db_xref="taxon:5664"
                     /chromosome="35"
                     /note="[illegible]"
BASE COUNT     8202 a  110[illegible] c  [illegible]  [illegible] others
ORIGIN

  Query Match             [illegible];  Score [illegible];  DB [illegible];  Length 375[illegible];
  Best Local Similarity   [illegible];  Pred. No. [illegible];
  Matches   38;  Conservative  [illegible];  Mismatches  [illegible];  Indels  [illegible];  Gaps  [illegible];

Qy          2 cgaatctaaacatatag[illegible]tcctgcatc 5[illegible]
              [illegible]
Db       7465 CGCATCCAG[illegible]ATC 7519

RESULT 7
```

```
AC_026408/c
LOCUS       AC026408   113996 bp    DNA             HTG       29-APR-2001
DEFINITION  Homo sapiens chromosome 5, clone CTB-226D8, WORKING DRAFT SEQUENCE,
            25 unordered pieces.
ACCESSION   AC026408
VERSION     AC026408.5  GI:13699543
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 113996)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 113996)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-MAR-2001) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Apr 29, 2001 this sequence version replaced gi:10305167.
            -----------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov

            Project Information
            Center Project Name: 22B621
            Center clone name: CTB-226D8

            Summary Statistics
            Consensus quality: 76916 bases at least Q40
            Consensus quality: 113845 bases at least Q30
            Consensus quality: 105987 bases at least Q20
            Estimated insert size: 121000; pulse field gel estimation
            Estimated insert size: 112096; sum of contigs estimation
            quality coverage: 4.0 in Q20 bases; pulse field gel estimation
            quality coverage: 4.31 in Q20 bases; sum of contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 25 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1      1179: contig of 1179 bp in length.
            *     1180      1279: gap of unknown length
            *     1280      3727: contig of 2448 bp in length.
            *     3728      3827: gap of unknown length
            *     3828      5283: contig of 1456 bp in length.
            *     5284      5383: gap of unknown length
            *     5384      7397: contig of 2014 bp in length.
            *     7398      7494: gap of unknown length
            *     7495      9408: contig of 1914 bp in length.
            *     9409      9509: gap of unknown length
            *     9510     12290: contig of 2781 bp in length.
            *    12291     12390: gap of unknown length
            *    12391     14324: contig of 1934 bp in length.
            *    14325     14424: gap of unknown length
            *    14425     17762: contig of 3338 bp in length.
            *    17763     17862: gap of unknown length
            *    17863     21982: contig of 4120 bp in length.
            *    21983     22082: gap of unknown length
            *    22083     26956: contig of 4874 bp in length.
            *    26957     27056: gap of unknown length
            *    27057     30272: contig of 3216 bp in length.
            *    30273     30372: gap of unknown length
            *    30373     34491: contig of 4119 bp in length.
            *    34492     34591: gap of unknown length
            *    34592     39875: contig of 5284 bp in length.
            *    39876     39975: gap of unknown length
            *    39976     46888: contig of 6913 bp in length.
            *    46889     46988: gap of unknown length
            *    46989     54208: contig of 7220 bp in length.
            *    54209     54308: gap of unknown length
            *    54309     62369: contig of 8061 bp in length.
            *    62370     62469: gap of unknown length
            *    62470     72463: contig of 9994 bp in length.
            *    72464     72563: gap of unknown length
            *    72564     86329: contig of 13766 bp in length.
            *    86330     86429: gap of unknown length
            *    86430    100776: contig of 14347 bp in length.
            *   100777    100876: gap of unknown length
            *   100877    113996: contig of 13120 bp in length.
FEATURES             Location/Qualifiers
     source          1..113996
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTB-226D8"
                     /clone_lib="CalTech human BAC library D"
BASE COUNT    41956 a  25987 c  25623 g  30909 t     1901 others
ORIGIN


  Query Match             40.9%;  Score 27.8;  DB 2;  Length 113996;
  Best Local Similarity   65.1%;  Pred. No. 49;
  Matches   41;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       1 tcagaatctaaacatataacaatatgaccaatcaaccaaccctcctcaaaaactcttcctgcatcatcacaa 60
           ||  | | | |||  || ||| ||| |||| | |    1  |||| |||||||| ||||
Db   28299 TCTGAcAtccAcAcAcAtcAtcc tcc tcc tcAAtcAcctcctcctcctcc tcAcAcAcAAA 28240

Qy      61 aaa 63
           |||
Db   28239 GAG 28237


RESULT   8
AC010607/c
LOCUS       AC010607   135545 bp    DNA             HTG       20-APR-2001
DEFINITION  Homo sapiens chromosome 5, clone CTB-28D11, WORKING DRAFT SEQUENCE,
            7 unordered pieces.
ACCESSION   AC010607
VERSION     AC010607.5  GI:13699543
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 135545)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 135545)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-SEP-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Apr 29, 2001 this sequence version replaced gi:7710621.
            -----------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            ----------

            Project Information
            Center Project Name: 75c39
            Center clone name: CTB-28SKB 28D11
            ----------

            Summary Statistics
            Consensus quality: 124665 bases at least Q40
            Consensus quality: 134287 bases at least Q30
            Consensus quality: 132099 bases at least Q20
            Estimated insert size: 160000; pulse field gel estimation
            Estimated insert size: 134945; sum of contigs estimation
```

```
            Quality coverage:  8.28 in Q20 bases; pulse field gel estimation.
            Quality coverage:  9.81 in Q20 bases; sum of contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 7 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1492: contig of 1492 bp in length
            *    1493     1492: gap of unknown length
            *    1493     6715: contig of 5223 bp in length
            *    6716     6815: gap of unknown length
            *    6816    12445: contig of 5630 bp in length
            *   12446    12545: gap of unknown length
            *   12546    27826: contig of 15281 bp in length
            *   27827    27926: gap of unknown length
            *   27927    41033: contig of 13107 bp in length
            *   41034    41133: gap of unknown length
            *   41134    64897: contig of 23764 bp in length
            *   64898    64997: gap of unknown length
            *   64998   145545: contig of 70548 bp in length.
FEATURES             Location/Qualifiers
     source          1..145545
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome "5"
                     /clone="CTB-28D11"
                     /clone_lib "Caltech human BAC library B"
BASE COUNT    39208 a   41408 c   29434 g   34984 t     611 others
ORIGIN


  Query Match             40.9%;  Score 27.8;  DB 2;  Length 145545;
  Best Local Similarity   65.1%;  Pred. No. 47;
  Matches   41;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       1 tcqaatctaqacata'qq cataqqtqqq qqcqc ctcqaaqqct ttcctqcatcacqq 60
           ||  |  | |||   | ||| ||| ||| | |     |  |||| |||||||||  ||||
Db 112169 TCTGAGATGGACTCATTGCCCTGGCTGGCTGGCTGGCTGGCTGTTCCTGCAGCACGG 112106

Qy      61 qaq 63
           |||
Db 112105 GAG 112103


RESULT   9
AC023060/c
LOCUS       AC023060    166173 bp    DNA             HTG       18-JUL-2000
DEFINITION  Homo sapiens chromosome 5 clone RP11 162M11, WORKING DRAFT
            SEQUENCE, 11 ordered pieces.
ACCESSION   AC023060
VERSION     AC023060.4  GI:9256436
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 166174)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 166174)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (26 FEB 2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 18, 2000 this sequence version replaced gi:7711755.
            --- --Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            ------
            Project Information
            Center Project Name: RP11
            Center clone name: RP11 162M11
            ------
            Summary Statistics
            Consensus quality:  [illegible] bases : low  [illegible]
            Consensus quality:  [illegible] bases : low  [illegible]
            Consensus quality:  [illegible] bases : low  [illegible]
            Estimated insert size: [illegible] agarose pulse estimator
            Estimated insert size: [illegible] sum of contig lengths
            Quality coverage:  [illegible] in [illegible] bases; pulse field gel estimation
            Quality coverage:  [illegible] in [illegible] bases; sum of contig lengths estimation
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 ordered pieces. The pieces
            * are represented as runs of N, but the exact sizes of the pieces
            * is believed to be correct as well as the order and the sizes
            * of the gaps between them. The exact sizes that have
            * provided by the submitter.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *       1    19708: contig of [illegible] bp length
            *   19709    [illegible]: gap of [illegible] bp
            *   19804    47579: contig of [illegible] bp length
            *   47580    47679: gap of [illegible] bp
            *   47680    72440: contig of [illegible] bp length
            *   72441    72540: gap of unknown length
            *   72541    75240: contig of [illegible] bp length
            *   75241    75340: gap of unknown length
            *   75341    [illegible]: contig of [illegible] bp length
            *   88859    88958: gap of unknown length
            *   88959    [illegible]: contig of [illegible] bp length
            *   99874    99973: gap of unknown length
            *   99974   102772: contig of [illegible] bp length
            *  102773   102852: gap of unknown length
            *  102853   111705: contig of [illegible] bp length
            *  111706   111805: gap of unknown length
            *  111806   156784: contig of [illegible] bp length
            *  156785   156884: gap of unknown length
            *  156885   164897: contig of [illegible] bp length
            *  164898   164997: gap of unknown length
            *  164998   [illegible]: contig of [illegible] bp length
FEATURES             Location/Qualifiers
     source          1..166173
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome "5"
                     /clone="RP11-162M11"
                     /clone_lib "Roswell Park Cancer Institute (PRI) Human BAC Library"
BASE COUNT    56168 a   47636 c   47454 g   [illegible] others
ORIGIN


  Query Match             40.9%;  Score 27.8;  DB [illegible];  Length 166173;
  Best Local Similarity   [illegible]%;  Pred. No. [illegible];
  Matches   41;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       1 tcqaatctaqacata'qq cataqqtqqq qqcqc ctcqaaqqct ttcctqcatcacqq 60
           |  |  | |||   | ||| ||| ||| | |     |  |||| |||||||||  ||||
Db 126446 TCTGAGATGGACTCATTGCCCTGGCTGGCTGGCTGGCTGGCTGTTCCTGCAGCACGG 126389

Qy      61 qaq 63
           |||
Db 126388 GAG 126386


RESULT  10
AL136321
LOCUS       AL136321   170642 bp    DNA             HTG       04-APR-[illegible]
DEFINITION  Human DNA sequence from clone [illegible] on chromosome [illegible]
            Contains [illegible] and CpG [illegible] map of [illegible]
ACCESSION   [illegible]
```

```
VERSION     AL163213.5  GI:27501524
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 170643)
  AUTHORS   Kimberley,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (04 APR 2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     on May 14, 2002 this sequence version replaced gi:27298852.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep. This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
            Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr1
            This sequence was finished as follows unless otherwise noted: all
            regions were either double stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >
            40); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. RPCI 11E23 is from
            the library RPCI 11.1 constructed by the group of Pieter de Jong.
            For further details see
            http://www.bcgsc.ca/bcgsc_projects/human.htm
            VERSION pDA(Feb.)
            This sequence is the entire insert of clone RP11 11E23 The true
            left end of clone RP11 16L9 is at 142476 in this sequence.
FEATURES             Location/Qualifiers
     source          1..170643
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1p1.3-p2.1"
                     /clone="RP11 11E23"
                     /clone_lib="RPCI 11.1"
     misc_feature    16..565
                     /note="match: GSS: Em:AQ427004"
     repeat_region   681..699
                     /note="23 copies 3 mer aga 76% conserved"
     repeat_region   1092..1308
                     /note="AluJb repeat: matches 1..200 of consensus"
     repeat_region   3527..3614
                     /note="MIR repeat: matches 60..151 of consensus"
     repeat_region   3617..3689
                     /note="MIR repeat: matches 180..260 of consensus"
     misc_feature    3878..4304
                     /note="match: STS: Em:G18001
                     match: GSS: Em:AQ284198"
     repeat_region   3894..4030
                     /note="MIR repeat: matches 7..126 of consensus"
     repeat_region   4414..4567
                     /note="MIR repeat: matches 79..228 of consensus"
     repeat_region   4765..4921
                     /note="MIR3 repeat: matches 1..186 of consensus"
     repeat_region   5272..5407
                     /note="9 copies 4 mer agac 100% conserved"
     repeat_region   5496..5551
                     /note="L1MC4 repeat: matches 1..157 of consensus"
     repeat_region   6241..6737
                     /note="L1MC4 repeat: matches 2202..2418 of consensus"
     repeat_region   6858..7414
                     /note="L2 repeat: matches 804..1417 of consensus"
     repeat_region   7780..8079
                     /note="L2 repeat: matches 1840..2151 of consensus"
     repeat_region   8244..8656
                     /note="L2 repeat: matches 2385..2497 of consensus"
     repeat_region   8519..8648
                     /note="16 copies 8 mer aaaaqaaa 63% conserved"
     repeat_region   8836..9248
                     /note="L1M4 repeat: matches 4826..5272 of consensus"
     misc_feature    complement(9371..9799)
                     /note="match: GSS: Em:AQ781278"
     misc_feature    complement(9383..9777)
                     /note="match: GSS: Em:AQ206508"
     misc_feature    complement(9491..9795)
                     /note="match: GSS: Em:AQ581541"
     misc_feature    9394..9815
                     /note="match: GSS: Em:B32946"
     misc_feature    complement(9468..9799)
                     /note="match: GSS: Em:AQ207860"
     misc_feature    9815..10291
                     /note="match: GSS: Em:AQ425540"
     repeat_region   9859..9884
                     /note="6 copies 4 mer gtgt 91% conserved"
     repeat_region   9862..9905
                     /note="4 copies 8 mer tgtgtgtg 94% conserved"
     misc_feature    9898..10330
                     /note="match: GSS: Em:AQ759451"
     repeat_region   9910..10274
                     /note="MER7A repeat: matches 1..346 of consensus"
     repeat_region   10016..10084
                     /note="AluJp repeat: matches 1..306 of consensus"
     repeat_region   11166..11280
                     /note="50 copies 2 mer aa 71% conserved"
     repeat_region   11816..11959
                     /note="MER1B repeat: matches 95..208 of consensus"
     misc_feature    12055..12447
                     /note="match: GSS: Em:AQ317924"
     misc_feature    12056..12529
                     /note="match: GSS: Em:AQ332817"
     repeat_region   12840..12974
                     /note="27 copies 5 mer aaatt 62% conserved"
     repeat_region   13131..13285
                     /note="L2 repeat: matches 2611..2710 of consensus"
     misc_feature    13264..13665
                     /note="match: GSS: Em:B40545
                     match: STS: Em:G37749"
     repeat_region   13557..13656
                     /note="MLT1I repeat: matches 1..308 of consensus"
     repeat_region   13747..13884
                     /note="L2 repeat: matches 2355..2496 of consensus"
     repeat_region   13925..13985
                     /note="L2 repeat: matches 2643..2709 of consensus"
     repeat_region   15042..15252
                     /note="MIR repeat: matches 1..252 of consensus"
     repeat_region   17808..18584
                     /note="MER21B repeat: matches 6..736 of consensus"
     repeat_region   19043..19251
                     /note="MER5A repeat: matches 1..187 of consensus"
     repeat_region   19433..19512
                     /note="10 copies 8 mer aaccaact 85% conserved"
     misc_feature    complement(20045..20493)
                     /note="match: STS: Em:AQ463010"
     repeat_region   23385..23728
                     /note="THE1C repeat: matches 1..370 of consensus"
     repeat_region   23729..25510
                     /note="THE1C internal repeat: matches 4..1580 of
                     consensus"
     repeat_region   25814..25870
                     /note="MSTA repeat: matches 1..426 of consensus"
     repeat_region   25824..26210
                     /note="MST repeat: matches 1..405 of consensus"
```

repeat_region 26264..26502
/note="MER20 repeat: matches 4. .238 of consensus"
repeat_region 26762..26813
/note="26 copies 2 mer tt 78% conserved"
repeat_region 28632..28820
/note="L1ME5 repeat: matches 274. .668 of consensus"
repeat_region 28815..29086
/note="MER30 repeat: matches 4. .189 of consensus"
misc_feature complement(29497..29907)
/note="match: GSS: Em:AQ022852"
repeat_region 29557..29600
/note="17 copies 4 mer ttca 81% conserved"
misc_feature 29680..29873
/note="match: STS: Em:G43867"
repeat_region 30511..30631
/note="MIR repeat: matches 47. .167 of consensus"
repeat_region 30437..31018
/note="MIR repeat: matches 5. .192 of consensus"
misc_feature complement(32337..32620)
/note="match: GSS: Em:AQ211265"
repeat_region 32611..32993
/note="MLI1B repeat: matches 1. .390 of consensus"
misc_feature 32626..33078
/note="match: GSS: Em:AQ268926"
repeat_region 33111..33163
/note="L2 repeat: matches 2635. .2682 of consensus"
repeat_region 33324..33363
/note="MIR repeat: matches 224. .262 of consensus"
repeat_region 33364..33661
/note="AluSc repeat: matches 3. .300 of consensus"
repeat_region 33762..33812
/note="MIR repeat: matches 67. .224 of consensus"
misc_feature 34256..34401
/note="match: GSS: Em:AQ581639"
repeat_region 34180..34483
/note="AluSp repeat: matches 4. .307 of consensus"
repeat_region 35438..35451
/note="MER2 repeat: matches 269. .343 of consensus"
repeat_region 35486..35697
/note="AluJb repeat: matches 82. .298 of consensus"
repeat_region 35698..35889
/note="MER2 repeat: matches 14. .207 of consensus"
repeat_region 36055..36384
/note="MLI1I repeat: matches 56. .410 of consensus"
repeat_region 36649..36726
/note="MIR repeat: matches 65. .141 of consensus"
repeat_region 38003..38482
/note="L2 repeat: matches 2033. .2458 of consensus"
repeat_region 38473..38568
/note="L2 repeat: matches 2604. .2699 of consensus"

Query Match 40.9%; Score 27.8; DB 9; Length 170642;
Best Local Similarity 69.1%; Pred. No. 44;
Matches 38; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 12 acatatgacgcat 24
||| ||||| |||||
Db 101437 ACAGATGGTCATTGGTGTACACAGCCAGGAGAGGCCTATGAGCATCAAGGTACAT 101491

Search completed: March 14, 2002, 15:10:43
Job time: 2508 sec